

# Binary Log-Linear Learning with Stochastic Communication Links

Arjun Muralidharan, Yuan Yan and Yasamin Mostofi

**Abstract**—In this paper, we consider distributed decision-making over stochastic communication links in multi-agent systems. We show how to extend the current literature on potential games with binary log-linear learning (which mainly focuses on ideal communication links) to consider the impact of stochastic communication channels. More specifically, we derive conditions on the probability of link connectivity to achieve a target probability for the set of potential maximizers (in the stationary distribution). Furthermore, our toy example demonstrates a transition phenomenon for achieving any target probability for the set of potential maximizers.

## I. INTRODUCTION

Non-cooperative game theory has recently emerged as a powerful tool for the distributed control of multi-agent systems [1]–[3]. By designing proper local utility functions and learning algorithms that satisfy certain properties, desirable global behaviors can be achieved. Potential games [4] are an important class of non-cooperative games and have recently received considerable attention in the literature [5]. In potential games, the local utility function of the agents is aligned with a potential function in order to achieve a global objective through local decisions.

There are a number of learning algorithms that can guarantee the convergence to a Nash equilibrium for potential games such as fictitious play [6] and joint strategy fictitious play [7]. However, a Nash equilibrium may be a sub-optimum outcome and not the potential maximizer. Log-linear learning (first introduced in [8]), on the other hand, is a learning mechanism that can guarantee convergence to the set of potential maximizers. As a result, it has been the subject of considerable research recently [9]. Binary log-linear learning [9], [10] is a variant of log-linear learning which can further handle constrained actions sets, i.e. scenarios where the future actions of the players are limited based on their current action (like in robotic networks).

While considerable progress has been made for distributed decision making using potential games, ideal communication links are often assumed. In other words, it is typically assumed that an agent can hear from all the other agents that will impact its utility function. In realistic communication environments with packet-dropping stochastic communication links, this is simply not possible. For instance, Fig. 1 shows an example of real channel measurements. We can see that the channel exhibits a great degree of stochasticity due to the shadowing and multipath fading components. Thus, it is the goal of this paper to bring an understanding of the impact of

stochastic packet-dropping communication links on potential games with binary log-linear learning, where each link is properly represented with an action-dependent probability of connectivity. By extending [9], we derive conditions on the temperature (defined in Section II-B) and probabilities of connectivity to achieve a given target probability (in the stationary distribution) for the set of potential maximizers (Theorems 3 and 4). In Section III-B, in a toy example, we further observe a transition behavior for achieving any target probability.

## II. PROBLEM SETUP

In this section, we first introduce some basic concepts and properties of potential games. We then review the binary log-linear learning algorithm and the theory of resistance trees, which we use in our subsequent analysis. Finally, we motivate the need for considering stochastic communication links.

### A. Potential Game (see [11] for more details)

A game  $\mathcal{G} = \{\mathcal{I}, \{\mathcal{A}_i\}_{i \in \mathcal{I}}, \{U_i\}_{i \in \mathcal{I}}\}$  is defined by its three components:

- 1)  $\mathcal{I} = \{1, 2, \dots, n\}$  is the set of players/agents/robots;
- 2)  $\mathcal{A}_i$  is the set of all the actions (choices) that agent  $i$  has. Then, an action profile  $a = (a_1, \dots, a_n) \in \mathcal{A}$  denotes the collection of actions of all the agents, where  $\mathcal{A} = \mathcal{A}_1 \times \dots \times \mathcal{A}_n$  is the space of all action profiles;
- 3)  $U_i : \mathcal{A} \rightarrow \mathbb{R}$  is the utility function of agent  $i$ .

One of the most important concepts in game theory is that of a pure Nash equilibrium, which is defined as follows.

**Definition 1 (Pure Nash Equilibrium):** Consider a game  $\mathcal{G} = \{\mathcal{I}, \{\mathcal{A}_i\}_{i \in \mathcal{I}}, \{U_i\}_{i \in \mathcal{I}}\}$ . An action profile  $a^{\text{NE}}$  is said to be a pure Nash equilibrium of the game if and only if

$$U_i(a^{\text{NE}}) \geq U_i(a_i, a_{-i}^{\text{NE}}), \quad \forall a_i \in \mathcal{A}_i \text{ and } \forall i \in \mathcal{I}, \quad (1)$$

where  $a_{-i} = (a_1, \dots, a_{i-1}, a_{i+1}, \dots, a_n) \in \mathcal{A}_{-i}$  denotes the action profile of all the agents except  $i$  and  $\mathcal{A}_{-i} = \mathcal{A}_1 \times \dots \times \mathcal{A}_{i-1} \times \mathcal{A}_{i+1} \times \dots \times \mathcal{A}_n$ .

As can be seen, a game has reached a pure Nash equilibrium if and only if no agent has the motivation to unilaterally change its action. In this paper, we are interested in potential games. Potential games can have broad applications in distributed multi-robot systems since they allow each robot to make local decisions while a global objective (the potential function) is optimized.

This work is supported in part by NSF NeTS award # 1321171.

The authors are with the Department of Electrical and Computer Engineering, University of California Santa Barbara, Santa Barbara, CA 93106, USA email: {arjunm, yuanyan, ymostofi}@ece.ucsb.edu.

**Definition 2 (Potential Games [4]):** A game  $\mathcal{G} = \{\mathcal{I}, \{\mathcal{A}_i\}_{i \in \mathcal{I}}, \{U_i\}_{i \in \mathcal{I}}\}$  is said to be a potential game with potential function  $\phi : \mathcal{A} \rightarrow \mathbb{R}$  if

$$U_i(a'_i, a_{-i}) - U_i(a_i, a_{-i}) = \phi(a'_i, a_{-i}) - \phi(a_i, a_{-i}), \quad \forall a_i, a'_i \in \mathcal{A}_i, \forall a_{-i} \in \mathcal{A}_{-i} \text{ and } \forall i \in \mathcal{I}. \quad (2)$$

As can be seen, a potential game requires a perfect alignment between the potential function and the agents' local utility functions. It is straightforward to confirm that the action profile that maximizes the potential function is a pure Nash equilibrium. Hence, a pure Nash equilibrium is guaranteed to exist in potential games.

### B. Binary Log-Linear Learning (see [9])

In several scenarios, the set of possible actions that an agent can take is limited by its current action. For instance, in multi-robot systems, the next possible position of an agent is limited by its current position and velocity. Formally, we refer to this limited set as an agent's constrained action set i.e.,  $\mathcal{A}_i^{\text{cons}}(a_i) \subseteq \mathcal{A}_i$  is agent  $i$ 's constrained action set where  $a_i$  is its current action.

Binary Log-linear learning (BLLL) is a variant of Log-linear learning (as shown in [9]) which can handle constrained action sets. It is summarized as follows. At each time step  $t$ , an agent  $i \in \mathcal{I}$  is chosen randomly (uniformly) and is allowed to alter its action.<sup>1</sup> All the other agents repeat their previous actions, i.e.  $a_{-i}(t) = a_{-i}(t-1)$ . Agent  $i$  then plays according to the following strategy:

$$p_i^{a_i(t-1)}(t) = \frac{e^{\frac{1}{\tau} U_i(a(t-1))}}{e^{\frac{1}{\tau} U_i(a(t-1))} + e^{\frac{1}{\tau} U_i(\hat{a}_i, a_{-i}(t-1))}}, \quad (3)$$

$$p_i^{\hat{a}_i}(t) = \frac{e^{\frac{1}{\tau} U_i(\hat{a}_i, a_{-i}(t-1))}}{e^{\frac{1}{\tau} U_i(a(t-1))} + e^{\frac{1}{\tau} U_i(\hat{a}_i, a_{-i}(t-1))}}, \quad (4)$$

where  $\hat{a}_i$  is an action that is chosen uniformly from the constrained action set  $\mathcal{A}_i^{\text{cons}}(a_i(t-1))$ ,  $p_i^{a_i(t-1)}(t)$  is the probability of repeating its previous action,  $p_i^{\hat{a}_i}(t)$  is the probability of selecting action  $\hat{a}_i$ , and  $\tau > 0$  is the temperature.

Moreover, the constrained action sets should possess the following two properties:

**Definition 3 (Reachability):** For all  $i \in \mathcal{I}$  and any action pair  $a_i^0, a_i^m \in \mathcal{A}_i$ , there exists a sequence of actions  $a_i^0 \rightarrow a_i^1 \rightarrow \dots \rightarrow a_i^{m-1} \rightarrow a_i^m$  satisfying  $a_i^k \in \mathcal{A}_i^{\text{cons}}(a_i^{k-1})$ ,  $\forall k \in \{1, \dots, m\}$ .

**Definition 4 (Reversibility):** For all  $i \in \mathcal{I}$  and any action pair  $a_i^0, a_i^1 \in \mathcal{A}_i$ , if  $a_i^1 \in \mathcal{A}_i^{\text{cons}}(a_i^0)$ , then we have  $a_i^0 \in \mathcal{A}_i^{\text{cons}}(a_i^1)$ .

Note that Definition 3 implies that any action profile in  $\mathcal{A}$  can be reached in finite time steps. Definition 4 means that each agent can go back to its previous action.

**Theorem 1:** (see [9]) Consider a potential game with constrained action sets that satisfy the reachability and reversibility properties. BLLL ensures that the support of the stationary distribution is the set of potential maximizers, as  $\tau \rightarrow 0$ .

<sup>1</sup>Note that the selection does not require coordination among the nodes and can be achieved through each agent using a Poisson clock [12].

We next introduce the concept of an asynchronous best reply process over constrained action sets, which is a process where each agent locally improves its own utility function when it is its turn to alter its action.

**Definition 5:** An asynchronous best reply process over constrained action sets is defined as follows. At each time  $t > 0$ , an agent  $i$  is randomly chosen (uniformly) and allowed to alter its action. All other agents repeat their current action, i.e.  $a_{-i}(t) = a_{-i}(t-1)$ . Agent  $i$  then selects an action  $\hat{a}_i$  uniformly from its constrained action set, i.e.  $\hat{a}_i \sim \text{unif}(\mathcal{A}_i^{\text{cons}}(a_i(t-1)))$ . It then plays the action which maximizes its utility function:  $a_i(t) \in \{a_i \in \{a_i, a_i(t-1)\} : U_i(a_i, a_{-i}(t-1)) = \max \{U_i(a(t-1)), U_i(\hat{a}_i, a_{-i}(t-1))\}\}$ .

The best reply process does not necessarily maximize the overall potential function of the game as it may result in a suboptimal Nash equilibrium. When  $\tau = 0$ , the BLLL algorithm boils down to an asynchronous best reply process on the constrained action sets. A  $\tau > 0$  then allows each agent to occasionally select locally suboptimal moves, i.e. it selects an action that decreases its local utility with a non-zero probability. These occasional suboptimal moves are useful as they prevent the agents from converging to a suboptimal Nash equilibrium. The BLLL algorithm can then be thought of as a perturbation of the asynchronous best reply process, where the size of the perturbation is controlled by the temperature  $\tau$ . This idea is formalized in Section II-C.2.

### C. Resistance Trees

In this part, we briefly review the concept of resistance trees, which we will use in our subsequent sections. We refer the readers to [13] for a detailed discussion.

1) **Resistance Trees:** Let  $P^0$  be a stationary Markov chain defined on a state space  $X$ . We call this the *unperturbed* process. The process  $P^\epsilon$  is then called a *regular perturbed Markov process* if it satisfies the following conditions:

- 1)  $P^\epsilon$  is aperiodic and irreducible;
- 2)  $\lim_{\epsilon \rightarrow 0} P^\epsilon(x \rightarrow y) = P^0(x \rightarrow y)$ ,  $\forall x, y \in X$ , where  $P^\epsilon(x \rightarrow y)$  and  $P^0(x \rightarrow y)$  denote the transition probabilities from state  $x$  to  $y$  of processes  $P^\epsilon$  and  $P^0$  respectively;
- 3) if  $P^\epsilon(x \rightarrow y) > 0$ , for some  $\epsilon > 0$ , then there exists some  $R(x \rightarrow y) \geq 0$ , such that  $0 < \lim_{\epsilon \rightarrow 0} \epsilon^{-R(x \rightarrow y)} P^\epsilon(x \rightarrow y) < \infty$ , where we refer to  $R(x \rightarrow y)$  as the resistance of the transition from state  $x$  to  $y$ .

Construct a tree  $T$  with  $|X|$  vertices, one for each state, rooted at some vertex  $z$ , such that there exists a unique directed path to  $z$  from every other vertex. The weight of a directed edge from vertex  $x$  to  $y$  is given by the resistance  $R(x \rightarrow y)$ . Such a tree is called a resistance tree whose resistance is given by the sum of the  $|X| - 1$  edges that compose it. Since  $P^\epsilon$  is aperiodic and irreducible (the first condition of the regular perturbed Markov process), there exists a unique stationary distribution  $\mu^\epsilon$  for a given  $\epsilon$ . Define  $p_z^\epsilon = \sum_{T \in \mathcal{T}_z} \prod_{[x,y] \in T} P^\epsilon(x \rightarrow y)$ , where  $[x, y]$  is

the directed edge from vertex  $x$  to  $y$  and  $\mathcal{T}_z$  denotes the set of all the trees that are rooted at  $z$ . Then, we have

$$\mu_z^\epsilon = \frac{p_z^\epsilon}{\sum_{z' \in X} p_{z'}^\epsilon}, \quad (5)$$

where  $\mu_z^\epsilon$  denotes the probability of state  $z$  in the stationary distribution.

The stochastic potential of state  $z$  is then defined as the minimum resistance among all the trees that are rooted at  $z$ :

$$\gamma(z) = \min_{T \in \mathcal{T}_z} \sum_{[x,y] \in T} R(x \rightarrow y), \quad (6)$$

*Theorem 2:* (see [13]) Let  $P^\epsilon$  be a regular perturbed Markov process of  $P^0$  and  $\mu^\epsilon$  be its unique stationary distribution. Then

- 1)  $\lim_{\epsilon \rightarrow 0} \mu^\epsilon = \mu^0$  exists,<sup>2</sup> where  $\mu^0$  is a stationary distribution of  $P^0$ ;
- 2) We have  $\mu_x^0 > 0$  iff  $\gamma(x) \leq \gamma(y)$ ,  $\forall y \in X$ , where  $\mu_x^0$  denotes the probability of state  $x$  in the stationary distribution  $\mu^0$ .

Theorem 2 shows that the stochastically stable states (the support of the stationary distribution  $\mu^0$ ) are the states with the minimum stochastic potential, i.e.  $\mu_x^0 > 0$  if and only if  $x$  minimizes  $\gamma(x)$ .

Informally, resistance (of a transition) is a measure of how difficult that transition is. The greater the resistance, the more difficult (less likely) the transition. So the resistance of a tree rooted at state  $x$  is a measure of how difficult it is for other states to transit to  $x$ . Thus, a state with minimum stochastic potential is a state to which it is the easiest to get to (informally speaking) as compared to other states. We will utilize this metaphor of difficulty in Section III-B when explaining some of our results.

2) *BLLL as a Regular Perturbed Markov process:* BLLL algorithm induces a regular perturbed Markov process with the unperturbed process corresponding to the asynchronous best reply process defined in Section II-B [9]. The probability of a feasible transition  $a^0 \rightarrow a^1 = (a_i^1, a_{-i}^0)$  (where agent  $i$  alters its action and  $a^0, a^1 \in \mathcal{A}$ ) is then given by

$$P^\epsilon(a^0 \rightarrow a^1) = \frac{1}{n|\mathcal{A}_i^{\text{cons}}(a_i^0)|} \frac{\epsilon^{-U_i(a_i^1, a_{-i}^0)}}{\epsilon^{-U_i(a_i^1, a_{-i}^0)} + \epsilon^{-U_i(a_i^0, a_{-i}^0)}}, \quad (7)$$

where  $\epsilon = e^{-\frac{1}{\tau}}$ . As shown in [9], the resistance of this transition is as follows:

$$R(a^0 \rightarrow a^1) = V_i(a^0, a^1) - U_i(a^1), \quad (8)$$

where  $V_i(a^0, a^1) = \max\{U_i(a^0), U_i(a^1)\}$ .

Based on the theory of resistance trees, it can be shown that only the action profiles that maximize the potential function have the minimum stochastic potential [13]. This in turn means that the stochastically stable states of BLLL are the set of potential maximizers, as also stated in Theorem 1.

<sup>2</sup>The perturbations effectively select one of the stationary distributions of  $P^0$ .

#### D. Stochastic Communication Links

Most of the current research in the area of motion planning of multi-robot systems assumes over-simplified channel models. For instance, it is common to assume perfect links or links that are perfect within a certain radius. In reality, however, communication links are best modelled stochastically. More specifically, the received channel to noise ratio (CNR) is a multi-scale random process with three major components: distance-dependent path loss, shadowing and multipath fading [14]. See Fig. 1 for a real example.

In the current literature on potential games, it is assumed that each agent is connected to all the other agents that can impact its next step utility function for all the possible actions in its constrained set. If the wireless channel is modeled as a disk with a known radius, then this can be achieved by properly designing the constrained action set. However, in the case of realistic communication links, this is simply not the case. More specifically, it is not possible for every agent to truly evaluate its utility function as other agents with whom it cannot communicate may be influencing it. Thus, realistic communication links have a considerable implication for distributed decision making using potential games. It is the goal of this paper to bring an understanding of their impact on BLLL and derive sufficient conditions (on link quality and temperature) to guarantee a target probability for the set of potential maximizers (in the stationary distribution) in the presence of stochastic links.

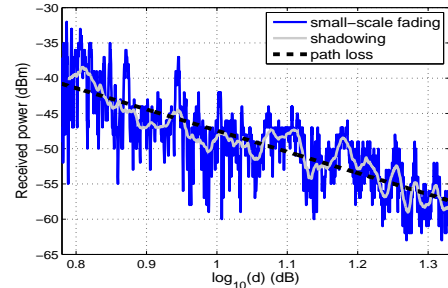


Fig. 1. Underlying dynamics of the received signal power across an indoor route [15].

### III. IMPACT OF STOCHASTIC COMMUNICATION LINKS ON BINARY LOG-LINEAR LEARNING

In this section, we characterize the impact of imperfect communication on the outcome of the BLLL algorithm. We first prove that given any arbitrarily-high probability of the set of potential maximizers, there exist a connectivity probability and temperature  $\tau$  that can achieve it. We then give an illustrative example to provide a deeper understanding of our results.

#### A. BLLL with Stochastic Communication Links

Consider the case where the communication graph among the agents is given by an undirected random graph  $\mathcal{C}(a) = (\mathcal{I}, \mathcal{E}(a))$ , where  $\mathcal{E}(a)$  denotes the set of edges, i.e. the communication links among the agents. Then, the probability of having a link (probability of connectivity) between agents

$i$  and  $j$  is given by  $p_{c,j,i}(a) = p_{c,i,j}(a) = p_c(a_i, a_j)$ , where we take  $p_{c,i,i}(a) = 1$ , for all  $a \in \mathcal{A}$  and  $i \in \mathcal{I}$ .<sup>3</sup> Note that we have taken the probability of connectivity (and subsequently the communication graph) to be action dependent to make our analysis more general (which naturally implies a time-varying graph). For instance, when the action of an agent involves its position, then the probability of connectivity becomes a function of the action profile. We assume that the probability of connectivity of different links are independent of each other in this paper. We further assume that the communication graph is drawn independently in each iteration.

As mentioned in Section II-B, in each iteration of the BLLL algorithm, an agent is chosen randomly (uniformly) to alter its action. Meanwhile, a realization of the communication graph is drawn from the random graph  $\mathcal{C}(a)$ . Let  $\mathcal{I}_{c,i}$  be the corresponding realization of the set of agents that agent  $i$  can communicate with. The probability of realization  $\mathcal{I}_{c,i}$  is given by  $p_{c,i}(\mathcal{I}_{c,i}, a) = \prod_{j \notin \mathcal{I}_{c,i}} (1 - p_{c,i,j}(a)) \prod_{j \in \mathcal{I}_{c,i}} p_{c,i,j}(a)$ . Note that  $\mathcal{I}_{c,i} = \mathcal{I}$  corresponds to the case where agent  $i$  can hear from all the other agents. Also, since the probability of connectivity is state-dependent,  $p_{c,i}(\mathcal{I}_{c,i}, a)$  is also a function of  $a$ .

The agent then has to assess its local utility and determine its action based on incomplete information. To represent this, we extend the definition of the utility function  $U_i : \mathcal{A} \rightarrow \mathbb{R}$  such that it is well defined for all  $U_i(a|\mathcal{I}_{c,i})$ ,  $\forall a \in \mathcal{A}, \forall \mathcal{I}_{c,i}$ , where  $U_i(a|\mathcal{I}_{c,i})$  is the evaluated local utility function of agent  $i$  given that it only communicates with agents in  $\mathcal{I}_{c,i}$ . One possibility for evaluating  $U_i(a|\mathcal{I}_{c,i})$  is that the agent ignores the impact of agents not in  $\mathcal{I}_{c,i}$ . Another possible strategy is for an agent to assume the last communicated action for the agents it is unable to communicate with.<sup>4</sup>

In order to evaluate the impact of the stochastic communication links on the learning dynamics, we start with a temperature-dependent probability of connectivity of the form  $p_{c,i,j}(a) = \frac{1}{1 + e^{m_{i,j}(a)}}$ ,  $\forall i, j \in \mathcal{I}, \forall a \in \mathcal{A}$ , where  $m_{i,j}(a) > 0$  is a constant. Based on our assumed form, we always have  $p_{c,i,j}(a) > 0.5$ . Note that for  $p_{c,i,j}(a) = p_c$ ,  $\forall a \in \mathcal{A}, \forall i, j \in \mathcal{I}$ ,  $p_c$  need not have this temperature-dependent form, as we will show in our result (Theorem 4). The probability of the transition  $a^0 \rightarrow a^1 = (a_i^1, a_{-i}^0)$  in the presence of stochastic communication links can then be characterized as follows:

$$\begin{aligned} P_c^\epsilon(a^0 \rightarrow a^1) &= \sum_{\mathcal{I}_{c,i}} p_{c,i}(\mathcal{I}_{c,i}, a^0) P_c^\epsilon(a^0 \rightarrow a^1 | \mathcal{I}_{c,i}) \\ &= \frac{1}{n |\mathcal{A}_i^{\text{cons}}(a_i^0)|} \sum_{\mathcal{I}_{c,i}} \frac{p_{c,i}(\mathcal{I}_{c,i}, a^0) \epsilon^{-U_i(a^1 | \mathcal{I}_{c,i})}}{\epsilon^{-U_i(a^1 | \mathcal{I}_{c,i})} + \epsilon^{-U_i(a^0 | \mathcal{I}_{c,i})}} \\ &= \frac{1}{n |\mathcal{A}_i^{\text{cons}}(a_i^0)|} \sum_{\mathcal{I}_{c,i}} \frac{\epsilon^{-U_i(a^1 | \mathcal{I}_{c,i}) + \sum_{j \notin \mathcal{I}_{c,i}} m_{i,j}(a^0)}}{(\epsilon^{-U_i(a^1 | \mathcal{I}_{c,i})} + \epsilon^{-U_i(a^0 | \mathcal{I}_{c,i})}) \prod_{j \in \mathcal{I}} (1 + e^{m_{i,j}(a^0)})}. \end{aligned}$$

<sup>3</sup>Notation not to be confused with  $p_z^\epsilon$ , which was used in (5).

<sup>4</sup>However, evaluating which is a better strategy becomes case dependent and is an avenue for future work.

It can be seen that expressing the probability of connectivity in this fashion ensures that BLLL in the presence of stochastic communication links induces a regular perturbed Markov process with the unperturbed process as the asynchronous best reply process (Definition 5).

We can further show that the equation above can be expressed as shown in (9) on top of the next page, which results in the following expression for the resistance of this transition:

$$R_c(a^0 \rightarrow a^1) = \min_{\mathcal{I}_{c,i}} \left\{ R_c(a^0 \rightarrow a^1 | \mathcal{I}_{c,i}) + \sum_{j \notin \mathcal{I}_{c,i}} m_{i,j}(a^0) \right\}, \quad (10)$$

where  $R_c(a^0 \rightarrow a^1 | \mathcal{I}_{c,i}) = V_i(a^0, a^1 | \mathcal{I}_{c,i}) - U_i(a^1 | \mathcal{I}_{c,i})$  and  $V_i(a^0, a^1 | \mathcal{I}_{c,i}) = \max\{U_i(a^0 | \mathcal{I}_{c,i}), U_i(a^1 | \mathcal{I}_{c,i})\}$ . Note that  $R_c(a^0 \rightarrow a^1 | \mathcal{I}) = R(a^0 \rightarrow a^1)$ , where  $R(a^0 \rightarrow a^1)$  is the resistance in case of perfect communication (see (8)).

It can be seen that imperfect communication affects the transition probability from  $a^0 \rightarrow a^1$ , and as a result, affects its resistance. This means that the stochastically stable states may change. Hence, the outcome of the game may be significantly different as compared to the case of perfect communication.

**Lemma 1:** Consider a potential game where the agents employ BLLL algorithm in the presence of stochastic communication links. Furthermore, consider constrained action sets that satisfy the reachability and reversibility properties. The states with the minimum stochastic potential are the set of potential maximizers if we have the following.

$$\sum_{j \notin \mathcal{I}_{c,i}} m_{i,j}(a^0) \geq R(a^0 \rightarrow a^1) - R_c(a^0 \rightarrow a^1 | \mathcal{I}_{c,i}), \quad (11)$$

for every agent  $i \in \mathcal{I}$ , all  $\mathcal{I}_{c,i}$  and all  $a^0 \rightarrow a^1 = (a_i^1, a_{-i}^0)$ , where  $R(a^0 \rightarrow a^1)$  is the resistance for the case of perfect communication (see (8)).

**Proof:** If the conditions in the lemma hold, then the resistance of the transition from  $a^0$  to  $a^1 = (a_i^1, a_{-i}^0)$ , for some agent  $i$ , becomes  $R_c(a^0 \rightarrow a^1) = R(a^0 \rightarrow a^1)$ . Therefore, the resistances of the transitions do not change as compared to the case of perfect communication. The proof of the lemma then follows immediately from Lemma 5.2 and Theorem 5.1 in [9]. ■

**Remark 1:** A good choice of the constants  $\{m_{i,j}(a)\}_{i,j \in \mathcal{I}, a \in \mathcal{A}}$ , is such that they satisfy  $m_{i,j}(a) \geq \max_{a^0 \rightarrow a^1} \{R(a^0 \rightarrow a^1)\}$ . This has the advantage that there are separate conditions for each  $m_{i,j}(a)$  and that they are not dependent on how communication failures affect the game.

**Remark 2:** Lemma 1 provides sufficient conditions to guarantee that the states with the minimum stochastic potential are still the set of potential maximizers. Equation (11) can be more explicitly expressed as a function of connectivity

$$P_c^\epsilon(a^0 \rightarrow a^1) = \frac{1}{n|\mathcal{A}_i^{\text{cons}}(a_i^0)|} \sum_{\mathcal{I}_{c,i}} \frac{\epsilon^{V_i(a^0, a^1|\mathcal{I}_{c,i}) - U_i(a^1|\mathcal{I}_{c,i}) + \sum_{j \notin \mathcal{I}_{c,i}} m_{i,j}(a^0)}}{(\epsilon^{V_i(a^0, a^1|\mathcal{I}_{c,i}) - U_i(a^1|\mathcal{I}_{c,i})} + \epsilon^{V_i(a^0, a^1|\mathcal{I}_{c,i}) - U_i(a^0|\mathcal{I}_{c,i})}) \prod_{j \in \mathcal{I}} (1 + \epsilon^{m_{i,j}(a^0)})}. \quad (9)$$

as follows:

$$\begin{aligned} \sum_{j \notin \mathcal{I}_{c,i}} m_{i,j}(a^0) &= \sum_{j \notin \mathcal{I}_{c,i}} \log_\epsilon(\epsilon^{m_{i,j}(a^0)}) \\ &= \sum_{j \notin \mathcal{I}_{c,i}} \log_\epsilon \frac{1 - p_{c,i,j}(a^0)}{p_{c,i,j}(a^0)} \\ &\geq R(a^0 \rightarrow a^1) - R_c(a^0 \rightarrow a^1|\mathcal{I}_{c,i}), \end{aligned}$$

for all  $\mathcal{I}_{c,i}$  and all  $a^0 \rightarrow a^1$ . We can see that  $\log_\epsilon \frac{1 - p_{c,i,j}(a^0)}{p_{c,i,j}(a^0)}$  is an important parameter (always positive). Furthermore, if  $R_c(a^0 \rightarrow a^1|\mathcal{I}_{c,i}) \geq R(a^0 \rightarrow a^1)$ , then connectivity to agent  $i$  is not important, since the condition is always satisfied.

The following theorem shows that we can find some  $\tau > 0$  and  $p_{c,i,j}(a) < 1$  to guarantee that the probability of the set of potential maximizers in the stationary distribution is larger than or equal to some required threshold.

**Theorem 3:** Consider a potential game where the agents employ BLLL algorithm in the presence of stochastic communication links. Furthermore, consider constrained action sets that satisfy the reachability and reversibility properties. For any given  $p_{\text{tar}} < 1$ , there exists a  $\tau_{\text{th}} > 0$ , such that the probability of the set of potential maximizers in the stationary distribution is larger than or equal to  $p_{\text{tar}}$ , if  $0 < \tau \leq \tau_{\text{th}}$  and  $p_{c,i,j}(a) = \frac{1}{1 + \epsilon^{m_{i,j}(a)}}$ ,  $\forall a \in \mathcal{A}$ ,  $\forall i, j \in \mathcal{I}$ , where  $\{m_{i,j}(a)\}_{i,j \in \mathcal{I}, a \in \mathcal{A}}$  are constants satisfying Lemma 1.

*Proof:* We construct temperature-dependent probabilities of connectivity, as discussed in Section III-A, such that the constants  $\{m_{i,j}(a)\}_{i,j \in \mathcal{I}, a \in \mathcal{A}}$  satisfy Lemma 1. Then, we have a regular perturbed Markov process, and the states with the minimum stochastic potential are still the set of potential maximizers.

From Theorem 2, we have  $\mu_x^0 = \lim_{\tau \rightarrow 0} \mu_x^\epsilon = 0$ ,  $\forall x \notin \mathcal{A}^*$ , where  $\mathcal{A}^* \subseteq \mathcal{A}$  is the set of potential maximizers. Thus, we know that, for any  $x \notin \mathcal{A}^*$ , there exists a  $\tau_x > 0$  such that  $\mu_x^\epsilon \leq \frac{1 - p_{\text{tar}}}{|\mathcal{A} \setminus \mathcal{A}^*|}$ , if  $0 < \tau \leq \tau_x$ . Hence, we have

$$\sum_{a \in \mathcal{A}^*} \mu_a^\epsilon = 1 - \sum_{a \notin \mathcal{A}^*} \mu_a^\epsilon \geq p_{\text{tar}},$$

if  $0 < \tau \leq \tau_{\text{th}} = \min_{x \notin \mathcal{A}^*} \tau_x$ . ■

The following theorem shows that we can find a sufficient lower bound on the probability of connectivity, to ensure that the probability of the set of potential maximizers is larger than or equal to some required threshold, for the special case of probabilities of connectivity which are state independent and equal for all links, i.e.,  $p_{c,i,j}(a) = p_c$ ,  $\forall a \in \mathcal{A}$ ,  $\forall i, j \in \mathcal{I}$ .

**Theorem 4:** Consider a potential game where the agents employ BLLL algorithm in the presence of stochastic communication links with probabilities of connectivity that are state independent and equal for all links, i.e.,  $p_{c,i,j}(a) = p_c$ ,  $\forall a \in \mathcal{A}$ ,  $\forall i, j \in \mathcal{I}$ . Furthermore, consider constrained action sets that satisfy the reachability and reversibility

properties. For any given  $p_{\text{tar}} < 1$ , there exists a  $p_{c,\text{th}} < 1$ , such that the probability of the set of potential maximizers in the stationary distribution is larger than or equal to  $p_{\text{tar}}$ , if  $p_c \geq p_{c,\text{th}}$ , where  $\tau = \frac{-m}{\ln(\frac{1-p_c}{p_c})}$ , with  $m$  representing a constant that satisfies Lemma 1.

*Proof:* We construct a temperature-dependent probability of connectivity, as discussed in Section III-A, such that the constant  $m$  satisfies Lemma 1, in order to establish a  $p_{c,\text{th}}$ . From Theorem 3, we know that there exists a  $\tau_{\text{th}} > 0$ , such that probability of convergence to the set of potential maximizers is larger than or equal to  $p_{\text{tar}}$ , if  $0 < \tau \leq \tau_{\text{th}}$ .

Consider  $p_c \geq p_{c,\text{th}} = \frac{1}{1 + e^{\frac{-m}{\tau_{\text{th}}}}}$ . Then, the temperature associated with this probability of connectivity is  $\tau = \frac{-m}{\ln(\frac{1-p_c}{p_c})} = \tau_{\text{th}} \frac{\ln(\frac{1-p_{c,\text{th}}}{p_{c,\text{th}}})}{\ln(\frac{1-p_c}{p_c})} \leq \tau_{\text{th}}$ . Thus, the probability of the set of potential maximizers is larger than or equal to  $p_{\text{tar}}$ , if  $p_c \geq p_{c,\text{th}}$  with  $\tau = \frac{-m}{\ln(\frac{1-p_c}{p_c})}$ . ■

Recall that the BLLL algorithm allows perturbations from the asynchronous best reply process, so as to allow agents to intentionally choose locally sub-optimum actions with a small probability. However, in the imperfect communication case, the failure of communication links also causes the agents to unintentionally make sub-optimum decisions. Given some  $p_{\text{tar}}$ , our given sufficient conditions then aim to restrict how often the unintentional mistakes can be made. One observation in the proof of Theorem 3 is that in general,  $\tau_{\text{th}}$  and  $p_{c,i,j}(a)$  are decreasing and increasing functions of  $p_{\text{tar}}$  respectively, i.e. if higher  $p_{\text{tar}}$  is required, then smaller temperature and better connectivity are needed. This is because when  $p_{\text{tar}}$  is higher, then less perturbation from the asynchronous best reply process is allowed. Hence, the agents have to assess their local utilities more accurately, which requires better connectivity.

Theorem 3 can be easily extended to the case of Log-Linear Learning. We skip the details for brevity.

## B. An illustrative example

In this part, we provide an example to have a better understanding on the impact of imperfect communication and the results of Theorem 3. Consider a 2-agent game where the action sets of the players are given by  $\mathcal{A}_1 = \{T, B\}$  (top, bottom) and  $\mathcal{A}_2 = \{L, R\}$  (left, right). The utility function  $U_i : \mathcal{A} = \mathcal{A}_1 \times \mathcal{A}_2 \rightarrow \mathbb{R}$  is given by

$U_1$	L	R	$U_2$	L	R
T	1	3	T	1	2
B	3	1	B	4	1

This is a potential game with the following potential function  $\phi : \mathcal{A} \rightarrow \mathbb{R}$ :

	$U_1$	$U_2$	$\phi$
$a^1 = (B, R)$	1	1	1
$a^2 = (T, L)$	1	1	2
$a^3 = (T, R)$	3	2	3
$a^4 = (B, L)$	3	4	4

For the case where the two nodes cannot communicate, we take  $U_1(a|\{1\}) = \begin{cases} 3 & \text{if } a_1 = T \\ 1 & \text{if } a_1 = B \end{cases}$  and  $U_2(a|\{2\}) = \begin{cases} 1 & \text{if } a_2 = L \\ 2 & \text{if } a_2 = R \end{cases}$ .

We next find the stochastic potential for each action profile, i.e. the minimum resistance of the tree rooted at each action profile, by using (6). For the case of perfect communication, the (only) state with minimum stochastic potential is  $a^4$ . For the case of imperfect communication, we consider the scenario where the probability of connectivity is state-independent, i.e.,  $p_c = p_{c,1,2} = p_{c,2,1}$ . Let  $m = m_{1,2} = m_{2,1} = \log_\epsilon \frac{1-p_c}{p_c}$ . By constructing all the resistance trees and minimizing (6), we find that  $a^4$  still has the minimum stochastic potential if and only if  $m = \log_\epsilon \frac{1-p_c}{p_c} > 1$ . From (11) in Lemma 1, our derived sufficient condition for the probability of connectivity to guarantee that  $a^4$  still has the minimum stochastic potential can be found as  $m = \log_\epsilon \frac{1-p_c}{p_c} \geq 3$ .

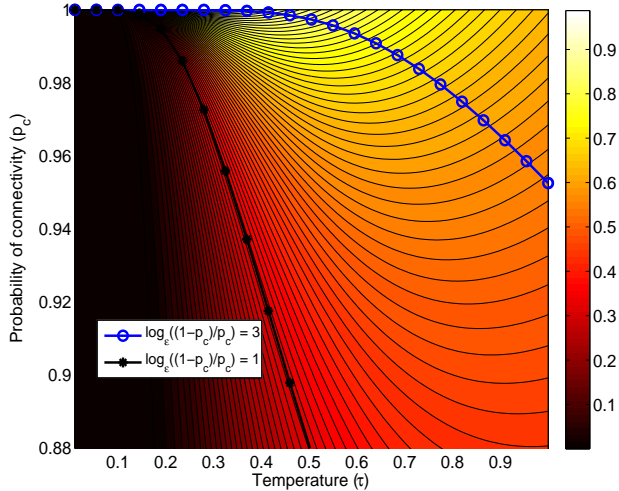


Fig. 2. The color map corresponds to the probability of the potential maximizer as a function of  $\tau$  and  $p_c$ . The blue line with empty circle markers shows  $p_c = \frac{1}{1+e^{-3/\tau}}$ , while the black line with filled circle markers shows  $p_c = \frac{1}{1+e^{-1/\tau}}$ .

By evaluating (5), we calculate the stationary distribution for the case of imperfect communication to see how the probability of the potential maximizer ( $a^4$ ) changes as a function of  $p_c$  and  $\tau$ , as shown in Fig. 2. The blue line with empty circle markers in the figure represents the curve  $p_c = \frac{1}{1+e^{-3/\tau}}$  ( $m = 3$ ), while the black line with filled circle markers represents  $p_c = \frac{1}{1+e^{-1/\tau}}$  ( $m = 1$ ). It can be seen that given any  $p_{tar}$ , the required probability of the potential maximizer can always be achieved by choosing a fixed  $m > 1$  and finding some appropriate  $\tau$  and  $p_c$ . (Theorem 3 shows a sufficient condition for this, where  $m \geq 3$ ).

Informally, as discussed in Section II-C, this is because the state with the minimum stochastic potential (in this case, the potential maximizer  $a^4$ ) is the easiest to transit to. The curve  $m = 1$ , i.e.,  $p_c = \frac{1}{1+e^{-1/\tau}}$  could be thought of as a transition curve, as curves above it can achieve any  $p_{tar}$ , while those below it cannot. This is due to the fact that above this curve  $a^4$  is the only state with minimum stochastic potential, while, on this curve, the states with minimum stochastic potential are  $a^4$  and  $a^3$ . On the other hand, below this curve,  $a^3$  is the (only) state with minimum stochastic potential. Informally, this means that the potential maximizer,  $a^4$ , becomes more difficult to transit to as compared to  $a^3$  when  $m < 1$ . In fact, for  $m < 1$ , the probability of the potential maximizer  $a^4$  becomes arbitrarily small as  $\tau \rightarrow 0$ . Note that we have plotted the y-axis only down to 0.88 for better visibility. Also, note that a case of connectivity of 96%, for instance, means that the packets are dropped 4% of the time, which is a typical value for several scenarios [16].

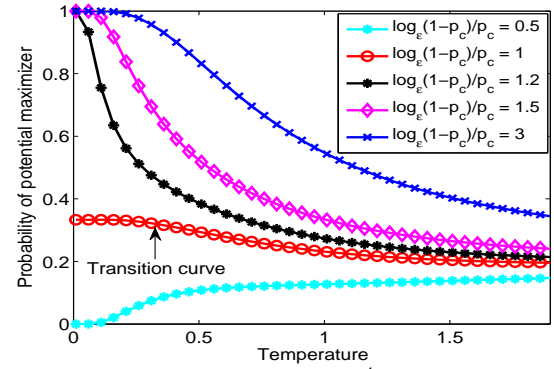


Fig. 3. Probability of potential maximizer ( $a^4$ ) as a function of temperature for different values of  $\log_\epsilon \frac{1-p_c}{p_c}$ .

Fig. 3 and 4 better highlight the transition behavior. Fig. 3 shows the probability of the potential maximizer as a function of the temperature, for various values of  $\log_\epsilon \frac{1-p_c}{p_c}$  ( $m$ ). The transition behavior of the curve  $m = \log_\epsilon \frac{1-p_c}{p_c} = 1$ , is clearly observed. Finally, Fig. 4 shows the probability of the potential maximizer as a function of  $\log_\epsilon \frac{1-p_c}{p_c}$  for various values of temperature  $\tau$ . The transition point can clearly be seen at  $\log_\epsilon \frac{1-p_c}{p_c} = 1$ .

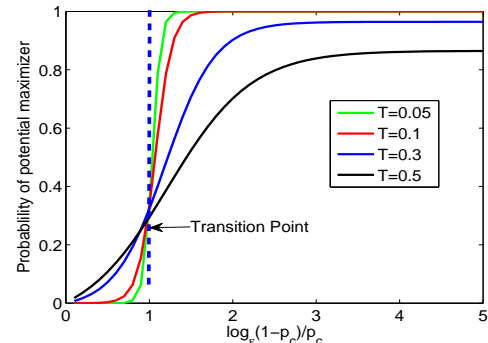


Fig. 4. Probability of potential maximizer ( $a^4$ ) vs  $\log_\epsilon \frac{1-p_c}{p_c}$  for different temperatures.

#### IV. CONCLUSIONS

In this paper, we considered the problem of distributed decision-making in multi-agent systems (via potential games) with an emphasis on the impact of realistic communication links. We showed how to extend the current literature on potential games with binary log-linear learning to account for stochastic communication channels. We derived conditions on the probabilities of link connectivity and BLLL's temperature to achieve a target probability for the set of potential maximizers. Furthermore, our toy example demonstrated a transition phenomenon for achieving any target probability.

#### REFERENCES

- [1] Jason R Marden, Gürdal Arslan, and Jeff S Shamma. Cooperative control and potential games. *Systems, Man, and Cybernetics, Part B: Cybernetics, IEEE Transactions on*, 39(6):1393–1407, 2009.
- [2] Minghui Zhu and Sonia Martínez. Distributed coverage games for mobile visual sensors (ii): Reaching the set of global optima. In *Decision and Control, 2009 held jointly with the 2009 28th Chinese Control Conference. CDC/CCC 2009. Proceedings of the 48th IEEE Conference on*, pages 175–180. IEEE, 2009.
- [3] Ahmet Yasin Yazicioglu, Magnus Egerstedt, and Jeff S Shamma. A game theoretic approach to distributed coverage of graphs by heterogeneous mobile agents. In *Estimation and Control of Networked Systems*, volume 4, pages 309–315, 2013.
- [4] Dov Monderer and Lloyd S Shapley. Potential games. *Games and economic behavior*, 14(1):124–143, 1996.
- [5] J. R. Marden and A. Wierman. Distributed welfare games. *Operations Research*, 61:155–168, 2013.
- [6] Dov Monderer and Lloyd S Shapley. Fictitious play property for games with identical interests. *Journal of economic theory*, 68(1):258–265, 1996.
- [7] Jason R Marden, Gürdal Arslan, and Jeff S Shamma. Joint strategy fictitious play with inertia for potential games. *Automatic Control, IEEE Transactions on*, 54(2):208–220, 2009.
- [8] Lawrence E Blume. The statistical mechanics of strategic interaction. *Games and economic behavior*, 5(3):387–424, 1993.
- [9] Jason Marden and Jeff S Shamma. Revisiting log-linear learning: Asynchrony, completeness and payoff-based implementation. *Games and Economic Behavior*, 75(2):788–808, 2012.
- [10] Gürdal Arslan, Jason R Marden, and Jeff S Shamma. Autonomous vehicle-target assignment: A game-theoretical formulation. *Journal of Dynamic Systems, Measurement, and Control*, 129(5):584–596, 2007.
- [11] Drew Fudenberg and Jean Tirole. *Game Theory*. MA: MIT Press, 1991.
- [12] Stephen Boyd, Arpita Ghosh, Balaji Prabhakar, and Devavrat Shah. Randomized gossip algorithms. *Information Theory, IEEE Transactions on*, 52(6):2508–2530, 2006.
- [13] H Peyton Young. The evolution of conventions. *Econometrica: Journal of the Econometric Society*, pages 57–84, 1993.
- [14] Andrea Goldsmith. *Wireless communications*. Cambridge university press, 2005.
- [15] Mehrzad Malmircheginia and Yasamin Mostofi. On the spatial predictability of communication channels. *Wireless Communications, IEEE Transactions on*, 11(3):964–978, 2012.
- [16] Jill M Boyce and Robert D Gaglianella. Packet loss effects on MPEG video sent over the public Internet. In *Proceedings of the sixth ACM international conference on Multimedia*, pages 181–190. ACM, 1998.